

<https://helda.helsinki.fi>

---

## A new approach to non-homogeneous local Tb theorems

Martikainen, Henri

2021-06

---

Martikainen , H , Mourgoglou , M & Vuorinen , E 2021 , ' A new approach to non-homogeneous local Tb theorems ' , Journal d'Analyse Mathematique , vol. 143 , no. 1 , pp. 95-121 . <https://doi.org/10.1007/s11854-021-0147-6>

---

<http://hdl.handle.net/10138/334387>

<https://doi.org/10.1007/s11854-021-0147-6>

---

acceptedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

# A NEW APPROACH TO NON-HOMOGENEOUS LOCAL $Tb$ THEOREMS

HENRI MARTIKAINEN, MIHALIS MOURGOLOU, AND EMIL VUORINEN

**ABSTRACT.** We develop a new general method to prove various non-doubling local  $Tb$  theorems. The method combines the non-homogeneous good lambda method of Tolsa, the big pieces  $Tb$  theorem of Nazarov–Treil–Volberg and a new change of measure argument based on stopping time techniques. We also improve known results and discuss some further applications.

## 1. INTRODUCTION

By a *singular integral operator* (SIO) related to some Borel measure  $\mu$  in  $\mathbb{R}^n$  we understand a linear operator  $T$  for which we have for all nice and disjointly supported functions  $f_1, f_2$  that

$$\int_{\mathbb{R}^n} T f_1(x) f_2(x) d\mu(x) = \iint_{\mathbb{R}^n \times \mathbb{R}^n} K(x, y) f_1(y) f_2(x) d\mu(y) d\mu(x)$$

for a suitable kernel  $K$  defined for  $x \neq y$ . An example in  $\mathbb{R}$  is the Hilbert transform, where  $d\mu(x) = dx$  and  $K(x, y) = 1/(x - y)$ . In general, SIOs need not be bounded in  $L^2(\mu)$ : the assumed size properties of the kernel  $K$  are too weak for this and more subtle cancellation has to be accounted for. This additional cancellation can most conveniently be formulated using various testing conditions. Such theorems characterising the boundedness of SIOs via testing conditions are called  $Tb$  theorems. The rough idea is that the action of  $T$  needs to be studied only on some single non-degenerate function  $b$  (such as  $b = 1$ ), or in the case of local  $Tb$  theorems, on a suitable family of non-degenerate functions  $b_Q$  indexed by cubes  $Q$  (such as  $b_Q = 1_Q$ ). An example of such a testing condition is the local  $T1$  condition:

$$\sup_{\substack{Q \subset \mathbb{R}^n \\ Q \text{ is a cube}}} \left[ \frac{1}{\mu(Q)} \int_Q |T1_Q|^2 d\mu + \frac{1}{\mu(Q)} \int_Q |T^*1_Q|^2 d\mu \right] < \infty.$$

The theory is significantly more subtle if the underlying measure  $\mu$  is non-doubling: such theory is called non-homogeneous Calderón–Zygmund analysis. The usual assumption then is that  $\mu(B(x, r)) \leq Cr^m$  for some exponent  $m$ , and that the kernel estimates for  $K$  are tied to this  $m$  (see Section 3).

---

2010 *Mathematics Subject Classification.* 42B20.

*Key words and phrases.* non-homogeneous analysis, Local  $Tb$  theorems, singular integrals, square functions.

H.M. was supported by the Academy of Finland through the grant Multiparameter dyadic harmonic analysis and probabilistic methods. Research of M.M. was supported by the ERC grant 320501 of the European Research Council (FP7/2007-2013). M.M. was also supported by IKERBASQUE and partially supported by the grant MTM-2017-82160-C2-2-P of the Ministerio de Economía y Competitividad (Spain). H.M. and E.V. are members of the Finnish Centre of Excellence in Analysis and Dynamics Research.

In this paper we develop a new method to prove certain very rough (meaning that the test functions  $b_Q$  can have very low integrability) local  $Tb$  theorems, and prove new and very general results. The theorems are formulated in the non-homogeneous situation. However, the results are already new even in the Lebesgue measure case. The developed method is very convenient and modular, and we will refer to multiple further works where it has been successfully applied.

**History and context.** The history of  $Tb$  theorems is extremely wide. The starting point in the Lebesgue measure case is the famous  $T1$  theorem by David–Journé [8]. Our focus is, however, on the local variants that were first introduced by M. Christ [5]. They are flexible tools as the non-degeneracy (also referred to as accretivity) of a given test function  $b_Q$  is only assumed on its supporting cube  $Q$ , i.e.,  $|\int_Q b_Q d\mu| \gtrsim \mu(Q)$ . In a global  $Tb$  theorem a single function  $b$  has to satisfy this non-degeneracy in all cubes, and constructing such a  $b$  can be more difficult. Christ’s version works for doubling measures  $\mu$  (but not more general than that) and requires that the test functions themselves are nice,  $b_Q \in L^\infty(\mu)$ , and that  $Tb_Q$  satisfies a demanding testing condition,  $Tb_Q \in L^\infty(\mu)$ . Very significant efforts have been made by multiple authors to allow *both* rougher test functions and less demanding testing conditions. A parallel line of investigation has dealt with the corresponding results in the non-homogeneous situation – we get to this later.

In general, we actually require two families of test functions – one for  $T$  and one for the adjoint  $T^*$ . In what follows we will always assume that  $T$  is antisymmetric,  $T^* = -T$ , which makes the discussion easier to follow. However, the completely general case sometimes involves very real difficulties – but we are not concerned with them here.

We say that a function  $b_Q$  is an  $L^p(\mu)$ -admissible,  $p \in [1, \infty)$ , test function on a cube  $Q \subset \mathbb{R}^n$  (with constant  $B_1$ ), if

- (1)  $\text{spt } b_Q \subset Q$ ,  $\mu(Q) = \int_Q b_Q d\mu$  and
- (2)  $(\frac{1}{\mu(Q)} \int_Q |b_Q|^p d\mu)^{1/p} \leq B_1$ .

A long standing problem (even for the Lebesgue measure) asks whether the  $L^2$  boundedness of (an antisymmetric) SIO  $T$  follows if we are given  $p \in (1, \infty)$ , and for every cube  $Q$  an  $L^p(\mu)$ -admissible test function  $b_Q$  (with the same constant  $B_1$ ) so that

$$(1.1) \quad \int_Q |Tb_Q|^{p'} d\mu \leq B_2 \mu(Q).$$

Here  $1/p + 1/p' = 1$ . This is often referred to as Hofmann’s local  $Tb$  problem. For certain simpler model operators this type of local  $Tb$  theorem in the Lebesgue measure case appears in Auscher–Hofmann–Muscalu–Tao–Thiele [1]. The extension to the SIO case has turned out to be of extreme difficulty – particularly if  $p < 2$ . In fact, for  $p < 2$  this is still not known in general. Even if the testing condition on  $1_Q Tb_Q$  becomes more demanding here as  $p'$  grows, i.e., when  $p$  gets smaller, the lower integrability of the test functions  $b_Q$  is desired. Hytönen–Nazarov [12] showed in the Lebesgue measure case that the  $L^2$  boundedness follows from the *buffered* testing condition  $\int_{2Q} |Tb_Q|^{p'} dx \lesssim |Q|$  for any  $p \in (1, \infty)$ . A key thing in the Lebesgue measure case is that if  $p \geq 2$ , then the original testing conditions (with  $1_Q Tb_Q$ ) automatically imply the stronger buffered testing conditions (with  $1_{2Q} Tb_Q$ ) by Hardy’s inequality. Previous related results include Auscher–Yang [4], Auscher–Routin [3] and Hofmann [10]. In the follow-up paper by two of us and Tolsa [17] we use and build on the methods presented in this paper to fully

prove the difficult case  $p < 2$  in  $\mathbb{R}$ , and show that we can always allow at least some exponents  $p \in (2 - \epsilon, 2]$  even when working in  $\mathbb{R}^n$  with  $n > 1$ .

We turn to discuss the non-homogeneous aspects in more detail. Non-doubling theory has been pioneered by David [7], Nazarov–Treil–Volberg [25] and Tolsa (see e.g. the book [26]). Previously, it was strongly believed that the class of doubling measures was the right class for the theory of singular integrals. The non-homogeneous theory is of extreme importance in many geometric questions – to mention just one key result, see the proof of the semiadditivity of analytic capacity by Tolsa [26, 27]. Moreover, Nazarov–Treil–Volberg [25] developed their so-called dyadic-probabilistic methods to deal with the difficulties posed by general measures, but such methods have also been extremely influential in other instances. For example, they led to the dyadic representation theorems and the first solution of the  $A_2$  conjecture by Hytönen [11].

Nazarov–Treil–Volberg [22] proved the first local  $Tb$  theorem in the non-doubling situation – there  $b_Q \in L^\infty(\mu)$  and  $Tb_Q \in \text{BMO}_2(\mu)$  (understood to mean that for all cubes  $R$  we have  $\int_R |Tb_Q - \langle Tb_Q \rangle_R^\mu|^2 d\mu \leq C\mu(2R)$ , where  $\langle Tb_Q \rangle_R^\mu = \mu(R)^{-1} \int_R Tb_Q d\mu$ ). In [13] one of us and Lacey proved a non-homogeneous local  $Tb$  theorem in the case that  $b_Q, 1_Q Tb_Q \in L^2(\mu)$  (even for SIOs  $T$  that are not necessarily antisymmetric). The state-of-the-art in the antisymmetric case is the already mentioned paper [17].

**Square functions vs SIOs.** It is of interest to consider local  $Tb$  theorems also for *square functions* (SFs)  $V$  that are introduced below. Such theorems have been applied e.g. in multiple Kato square root papers – see for instance [2]. See also Example 1.5 below. For us, however, SFs mainly just offer a simpler platform compared to SIOs, and allows us to present the full technical execution of our method, including the big piece  $Tb$  type argument discussed in more detail below. A key technical difference between SFs  $V$  and SIOs  $T$  is that for SIOs our testing conditions need to a priori involve  $1_Q T_* b_Q$  (instead of  $1_Q Tb_Q$ ), where  $T_*$  is the maximally truncated SIO

$$T_* f(x) = \sup_{\epsilon > 0} |T_\epsilon f(x)|, \quad T_\epsilon f(x) = \int_{|x-y| > \epsilon} K(x, y) f(y) d\mu(y).$$

New local  $Tb$  theorems for antisymmetric Calderón–Zygmund operators with testing conditions involving  $1_Q T_* b_Q$  follow at once from the ideas of this paper, and we explicitly state these in Section 3. To pass from such results to the original Hofmann’s problem (with testing conditions involving  $1_Q Tb_Q$  as in (1.1)) a so-called adapted Cotlar’s inequality is needed: one needs to use the existence of the test functions to prove a Cotlar type inequality, i.e., a pointwise control of  $T_*$  by something involving only  $T$ . A fancy version of an adapted Cotlar’s inequality is presented in the follow-up paper [17] (for previous version see also [12]). This step causes the fact that in Hofmann’s problem we can only allow  $p \in (2 - \epsilon, 2]$  (that is, we cannot always go all the way down to  $p > 1$  in [17]).

Results involving  $V$  and  $T_*$ , however, behave in completely analogous ways. In such formulations we can allow  $L^p(\mu)$ -admissible test functions with any  $p > 1$ , and, as we shall soon see, we can do quite a bit more – we can e.g. even use test measures with weaker conditions than this. Moreover, the testing conditions need not be as in Hofmann’s problem (1.1). In fact, we can have  $1_Q V b_Q$  (or  $1_Q T_* b_Q$ ) in any  $L^{s, \infty}$ ,  $s > 0$ . That is, the given regularity of the test function (the exponent  $p$ ) need not be reflected in the testing condition at all (we can use any  $s > 0$ ) – the conditions decouple.

We now define the SFs that we use. Let  $m, \alpha > 0$ , and assume that we have kernels  $s_t: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{C}$ ,  $t > 0$ , satisfying the size condition

$$(1.2) \quad |s_t(x, y)| \lesssim \frac{t^\alpha}{(t + |x - y|)^{m+\alpha}},$$

the  $y$ -Hölder condition

$$(1.3) \quad |s_t(x, y) - s_t(x, z)| \lesssim \frac{|y - z|^\alpha}{(t + |x - y|)^{m+\alpha}}$$

whenever  $|y - z| < t/2$ , and the  $x$ -Hölder condition

$$(1.4) \quad |s_t(x, y) - s_t(z, y)| \lesssim \frac{|x - z|^\alpha}{(t + |x - y|)^{m+\alpha}}$$

whenever  $|x - z| < t/2$ . Let  $M(\mathbb{R}^n)$  denote the vector-space of all complex Borel measures in  $\mathbb{R}^n$ . The variation measure of  $\nu \in M(\mathbb{R}^n)$  is denoted by  $|\nu|$  and the total variation is  $\|\nu\| = |\nu|(\mathbb{R}^n)$ . For a given complex measure  $\nu$  we define

$$\theta_t \nu(x) = \int s_t(x, y) d\nu(y), \quad x \in \mathbb{R}^n.$$

The vertical square function  $V$  is defined by

$$V\nu(x) = \left( \int_0^\infty |\theta_t \nu(x)|^2 \frac{dt}{t} \right)^{1/2}, \quad x \in \mathbb{R}^n.$$

Given a cube  $Q \subset \mathbb{R}^n$  define also the truncated version

$$V_Q \nu(x) = \left( \int_0^{\ell(Q)} |\theta_t \nu(x)|^2 \frac{dt}{t} \right)^{1/2},$$

where  $\ell(Q)$  denotes the side length of  $Q$ .

We say that a (positive) Radon measure  $\mu$  in  $\mathbb{R}^n$  is of order  $m$ , if  $\mu(B(x, r)) \lesssim r^m$  for all  $x \in \mathbb{R}^n$  and  $r > 0$ . For  $f \in \bigcup_{p \in [1, \infty]} L^p(\mu)$  and  $x \in \mathbb{R}^n$  we set

$$\theta_t^\mu f(x) := \theta_t(f d\mu) = \int s_t(x, y) f(y) d\mu(y)$$

and

$$V_\mu f(x) := V(f d\mu)(x) = \left( \int_0^\infty |\theta_t^\mu f(x)|^2 \frac{dt}{t} \right)^{1/2}.$$

Define also  $V_{\mu, Q} f = V_Q(f d\mu)$ . The above definitions make sense also when  $\mu$  is finite (and not necessarily of order  $m$ ).

**1.5. Example.** Mayboroda–Volberg [21], Chousionis, Garnett, Le and Tolsa [6] and others have linked boundedness of square functions and geometry. Let  $E \subset \mathbb{R}^n$  be a closed set which is  $m$ -ADR for some integer  $0 < m < n$ , i.e.,  $\mu := \mathcal{H}^m|_E$  satisfies  $\mu(B(x, r)) \sim r^m$  for all  $x \in E$  and  $r \in (0, \text{diam}(E))$ . Let  $s_t(x, y) = t \partial_t [t^{-m} \phi([x - y]/t)]$ , where  $\phi(x) = (1 + |x|^2)^{-(m+1)/2}$ . This kernel satisfies (1.2), (1.3) and (1.4). One of the results of [6] says that  $E$  is uniformly  $m$ -rectifiable (for the definition see [6]) if and only if  $V_\mu$  is  $L^2(\mu)$  bounded.

**Description of the new method and applications.** In the first step of the method the idea is to fix a cube  $Q$  and prove that there exists  $G_Q \subset Q$  so that  $\mu(G_Q) \gtrsim \mu(Q)$  and  $\|1_{G_Q} V_\mu f\|_{L^2(\mu)} \lesssim \|f\|_{L^2(\mu)}$  for every  $f \in L^2(\mu)$  satisfying that  $\text{spt } f \subset G_Q$ . This relies on the big pieces type  $Tb$  theorem of Nazarov, Treil and Volberg [24] – see also Volberg’s book [28] and Tolosa’s book [26] for expositions of this theorem. In the appendix we also formulate and give a complete proof of the big piece  $Tb$  theorem in the SF situation, which is much more approachable than the difficult one in the SIO context (compare e.g. to the book of Tolosa [26], pages 137–194). We hope that our proof helps to make this important theorem more approachable. In our rough local  $Tb$  setting this theorem certainly cannot be applied directly – it e.g. requires a bounded test function. The idea is to first perform a change of measure to  $\sigma := |b_Q| d\mu$  and aim to apply the big piece  $Tb$  theorem A.1 with the measure  $\sigma$  and the  $L^\infty$ -function  $b_Q/|b_Q|$ . In essence, we make the measure worse but the function a lot better. In order to pass from the  $\mu$  measure to the  $\sigma$  measure, and then back, we use stopping time arguments.

The non-homogeneous good lambda method of Tolosa [26] – see Theorem 2.5 below – says that the existence of such a big piece  $G_Q$  in every nice cube (doubling and of small boundary) is enough to guarantee the global  $L^2(\mu)$  boundedness of  $V_\mu$ . It is extremely convenient that the good lambda method is so flexible – this is the reason why we can assume the existence of  $b_Q$  only in nice cubes.

This general method is the main contribution of this article. Previously one of the key difficulties in proving non-homogeneous local  $Tb$  theorems with rough test functions was that it was difficult to prove the boundedness of certain  $b_Q$ -adapted (or twisted) martingale transforms (see Lacey–Martikainen [13]). We do not need such transformations in our new strategy, and this is highly convenient.

We compare our method to that of Hytönen–Nazarov [12]. Their proof works for doubling measures and it does not appear to be clear how to extend their method to non-homogeneous measures. Moreover, they do not use the big pieces global  $Tb$  or the good lambda method. Instead, they write a rough test function as a sum of the good and bad part using the Calderón–Zygmund decomposition. The good part is a non-degenerate  $L^\infty$  function which they want to use as a test function in some simpler  $Tb$  theorem. To transfer their testing conditions involving  $Tb_Q$  to this new bounded function they have to perform stopping times in a delicate order, and suppress their operator appropriately in the bad set. Thus, their proof is very different from our strategy. In [17] we prove a deeper instance of Cotlar’s inequality than in [12], and then combine this with our local  $Tb$  theorems involving  $T_*$  (Proposition 3.1 below). This combination then allows us to drop their buffer assumption even for some  $p < 2$ , and gives a proof that works in the non-homogeneous situation.

Besides the best known local  $Tb$  theorem for SFs and maximally truncated SIOs  $T_*$  presented in this paper and the best known result concerning Hofmann’s problem [17], we have also applied this new method in [18], where the flexibility of the method is helpful in handling issues involving metric space arguments. Lastly, we mention that in [21] we applied the ideas of the current paper and [17] to prove new results in the bilinear setting.

**Formulation of the new local  $Tb$  theorem for square functions.** A cube  $Q \subset \mathbb{R}^n$  is called  $(\alpha, \beta)$ -doubling for a given measure  $\mu$  if  $\mu(\alpha Q) \leq \beta \mu(Q)$ . Given  $t > 0$  we say that a cube



$Q \subset \mathbb{R}^n$  has  $t$ -small boundary with respect to the measure  $\mu$  if

$$\mu(\{x \in 2Q: \text{dist}(x, \partial Q) \leq \lambda \ell(Q)\}) \leq t\lambda\mu(2Q)$$

for every  $\lambda > 0$ . The following theorem presents the best known local  $Tb$  theorem for SFs.

**1.6. Theorem.** *Let  $\mu$  be a measure of order  $m$  in  $\mathbb{R}^n$  and  $B_1, B_2 < \infty$ ,  $\epsilon_0 \in (0, 1)$  be given constants. Let  $\beta > 0$  and  $C_1$  be large enough (depending only on  $n$ ). Suppose that for every  $(2, \beta)$ -doubling cube  $Q \subset \mathbb{R}^n$  with  $C_1$ -small boundary there exists a complex measure  $\nu_Q$  so that*

- (1)  $\text{spt } \nu_Q \subset Q$ ;
- (2)  $\mu(Q) = \nu_Q(Q)$ ;
- (3)  $\|\nu_Q\| \leq B_1\mu(Q)$ ;
- (4) For all Borel sets  $A \subset Q$  satisfying  $\mu(A) \leq \epsilon_0\mu(Q)$  we have

$$|\nu_Q|(A) \leq \frac{\|\nu_Q\|}{32B_1}.$$

Suppose there exist  $s > 0$  and for all  $Q$  as above a Borel set  $U_Q \subset \mathbb{R}^n$  such that  $|\nu_Q|(U_Q) \leq \frac{\|\nu_Q\|}{16B_1}$  and

$$\sup_{\lambda > 0} \lambda^s \mu(\{x \in Q \setminus U_Q: V_Q \nu_Q(x) > \lambda\}) \leq B_2 \|\nu_Q\|.$$

Then  $V_\mu: L^p(\mu) \rightarrow L^p(\mu)$  for every  $p \in (1, \infty)$ .

The following points are aimed to clarify the technical aspects of this theorem.

- If  $V_\mu: L^2(\mu) \rightarrow L^2(\mu)$  boundedly, then  $V: M(\mathbb{R}^n) \rightarrow L^{1,\infty}(\mu)$  boundedly. In this non-homogeneous setting this is (for SIOs) a result of Nazarov–Treil–Volberg [23] – for another proof see [26]. It follows that given  $\nu_Q$  like above one has to have

$$\sup_{\lambda > 0} \lambda \mu(\{x \in Q: V_Q \nu_Q(x) > \lambda\}) \leq \sup_{\lambda > 0} \lambda \mu(\{x: V \nu_Q(x) > \lambda\}) \leq C \|\nu_Q\|.$$

This makes the assumptions necessary.

- Suppose  $q \in (1, \infty)$  and that  $b_Q$  is an  $L^q(\mu)$ -admissible test function with a constant  $B_1$ . Then  $\nu_Q := b_Q d\mu$  is a testing measure as in the above theorem – i.e., it satisfies (1)–(4). Here  $q = 1$  is enough for (3) but  $q > 1$  can be used to get (4):

$$\int_A |b_Q| d\mu \leq \mu(A)^{1/q'} B_1 \mu(Q)^{1/q} \leq \epsilon_0^{1/q'} B_1 \int_Q b_Q d\mu \leq \frac{1}{32B_1} \int_Q |b_Q| d\mu$$

if  $\epsilon_0 := (32B_1^2)^{-q'}$ . In general, it is enough to prove (4) by some method.

- It is to be expected that the fact that the cubes are doubling and of small boundary makes the theorem significantly easier to apply with non-homogeneous measures than the previously known theorems. We are not aware how to modify the strategies in [13], [14] and [16] to yield this generality.
- The fact that we can allow a small exceptional set  $U_Q$  allows some additional flexibility in technical arguments, and this is in fact exploited in [17].

In the previously known state-of-the-art theorems for SFs the assumptions have been of the form that  $q \in (1, \infty)$ , for each cube  $Q$  we are given an  $L^q(\mu)$ -admissible test function  $b_Q$  and the testing condition holds with the same  $q$ :

$$\int_Q |V_{\mu,Q} b_Q|^q d\mu \lesssim \mu(Q).$$

The Lebesgue case is by Hofmann [9] and the non-doubling case by two of us [16]. Previous results (the case  $q = 2$ ) include [2] and [14].

In Section 2 we prove our main result for SFs and present our method in detail. In Section 3 we discuss the analogous results for SIOs. In Appendix A we give a relatively short proof of the big pieces global  $Tb$  theorem for square functions, which we need to apply in Section 2.

**Notation.** We write  $A \lesssim B$ , if there is a constant  $C > 0$  so that  $A \leq CB$ . We may also write  $A \sim B$  if  $B \lesssim A \lesssim B$ .

We then set some dyadic notation. For cubes  $Q$  and  $R$  we denote

- $\ell(Q)$  is the side-length of  $Q$ ;
- $d(Q, R)$  denotes the distance between the cubes  $Q$  and  $R$ ;
- $D(Q, R) := d(Q, R) + \ell(Q) + \ell(R)$  is the long distance;
- $W_Q = Q \times [\ell(Q)/2, \ell(Q))$  is the Whitney region associated with  $Q$ ;
- $\text{ch}(Q)$  denotes the dyadic children of  $Q$ ;
- $\mu|_Q$  denotes the measure  $\mu$  restricted to  $Q$ ;
- $\langle f \rangle_Q^\mu = \mu(Q)^{-1} \int_Q f d\mu$  (or just  $\langle f \rangle_Q$  if the measure is clear from the context).

**Acknowledgements.** Much of this research was carried out when H.M. was visiting CRM during September 2015, and he would like to thank the institution for its hospitality. We thank Benjamin Jaye for answering our question regarding a technical point in some  $Tb$  theorems. We would also like to thank the anonymous referee and the editor for comments that helped to clarify the exposition.

## 2. THE MAIN METHOD AND PROOF OF THE LOCAL $Tb$ THEOREM FOR SFs

We record the following easy lemma.

**2.1. Lemma.** *Let a cube  $Q \subset \mathbb{R}^n$  be given and  $G \subset Q$ . Suppose also that  $\nu(Q) \lesssim \ell(Q)^m$ . If  $\|1_G V_{\nu, Q} f\|_{L^2(\nu)} \lesssim \|f\|_{L^2(\nu)}$  for every  $f \in L^2(\nu)$  satisfying  $\text{spt } f \subset G$ , then also  $\|1_G V_{\nu} f\|_{L^2(\nu)} \lesssim \|f\|_{L^2(\nu)}$  for every  $f \in L^2(\nu)$  satisfying  $\text{spt } f \subset G$ .*

*Proof.* This follows from the pointwise estimate

$$\int_{\ell(Q)}^{\infty} |\theta_t^\nu f(x)|^2 \frac{dt}{t} \lesssim \nu(Q) \ell(Q)^{-2m} \|f\|_{L^2(\nu)}^2.$$

□

**2.2. Definition.** Given a cube  $Q \subset \mathbb{R}^n$  we consider the following random dyadic grid. For small notational convenience assume that  $c_Q = 0$  (that is,  $Q$  is centred at the origin). Let  $N \in \mathbb{Z}$  be defined by the requirement  $2^{N-3} \leq \ell(Q) < 2^{N-2}$ . Consider the random square  $Q^* = Q^*(w) = w + [-2^N, 2^N]^n$ , where  $w \in [-2^{N-1}, 2^{N-1}]^n =: \Omega_N = \Omega$ . The set  $\Omega$  is equipped with the normalised Lebesgue measure  $\mathbb{P}_N = \mathbb{P}$ . We define the grid  $\mathcal{D}(w) := \mathcal{D}(Q^*(w))$ . Notice that  $Q \subset \alpha Q^*(w)$  for some  $\alpha < 1$ , and  $\ell(Q) \sim \ell(Q^*(w))$ .

Next, we prove the main Proposition.

**2.3. Proposition.** *Let  $\mu$  be a measure of order  $m$  and  $B_1, B_2 < \infty$ ,  $\epsilon_0 \in (0, 1)$  be given constants. Let  $Q \subset \mathbb{R}^n$  be a fixed cube. Assume that there exists a complex measure  $\nu = \nu_Q$  such that*

- (1)  $\text{spt } \nu \subset Q$ ;



- (2)  $\mu(Q) = \nu(Q)$ ;
- (3)  $\|\nu\| \leq B_1\mu(Q)$ ;
- (4) For all Borel sets  $A \subset Q$  satisfying  $\mu(A) \leq \epsilon_0\mu(Q)$  we have

$$|\nu|(A) \leq \frac{\|\nu\|}{32B_1}.$$

Suppose there exist  $s > 0$  and a Borel set  $U_Q \subset \mathbb{R}^n$  for which  $|\nu|(U_Q) \leq \frac{\|\nu\|}{16B_1}$  so that

$$\sup_{\lambda > 0} \lambda^s \mu(\{x \in Q \setminus U_Q : V_Q \nu(x) > \lambda\}) \leq B_2 \|\nu\|.$$

Then, there is some subset  $G_Q \subset Q \setminus U_Q$  such that  $\mu(G_Q) \gtrsim \mu(Q)$  and

$$\|1_{G_Q} V_\mu f\|_{L^2(\mu)} \lesssim \|f\|_{L^2(\mu)}$$

for every  $f \in L^2(\mu)$  satisfying that  $\text{spt } f \subset G_Q$ .

*Proof.* We can assume that  $\text{spt } \mu \subset Q$ . Indeed, if we have proved the theorem for such measures, we can then apply it to  $\mu|_Q$ . Let us denote  $\sigma = |\nu|$ , where  $|\nu|$  is the variation measure of  $\nu$ . Also, let us write the polar decomposition of the complex measure  $\nu$  as  $\nu = b d\sigma$ , where  $b$  is a function so that  $|b(x)| = 1$  always.

The idea is to apply the big pieces global  $Tb$  theorem from Appendix A (Theorem A.1). It will be applied to the measure  $\sigma$  and the bounded function  $b$ . Using stopping times we need to construct some exceptional sets so that the assumptions of that theorem are verified. Moreover, we need to be able to come back to the  $\mu$  measure – this requires encompassing additional stopping times to the construction.

We fix  $w$ , and write  $\mathcal{D}(w) = \mathcal{D}$ . We also write  $\mathcal{D}_0 = \mathcal{D}(0)$ . Let  $\mathcal{A} = \mathcal{A}_w$  consist of the maximal dyadic cubes  $R \in \mathcal{D}$  for which

$$\left| \int_R b d\sigma \right| < \eta \sigma(R),$$

where  $\eta := \frac{1}{2}B_1^{-1}$ . We set

$$T = T_w = \bigcup_{R \in \mathcal{A}} R \subset \mathbb{R}^n.$$

Notice that

$$\sigma(Q) = \|\nu\| \leq B_1\mu(Q) = B_1\nu(Q) = B_1 \int_Q b d\sigma.$$

Then estimate

$$\int_Q b d\sigma = \left| \int_Q b d\sigma \right| = \left| \int_{Q \setminus T} b d\sigma + \sum_{R \in \mathcal{A}} \int_R b d\sigma \right| \leq \sigma(Q \setminus T) + \eta \sigma(Q).$$

Since  $\eta B_1 = 1/2$  we conclude that

$$\sigma(Q) \leq B_1 \sigma(Q \setminus T) + \frac{1}{2} \sigma(Q),$$

and so

$$\sigma(Q) \leq 2B_1[\sigma(Q) - \sigma(T)].$$

From here we can read that

$$\sigma(T) \leq (1 - \eta)\sigma(Q).$$

Next, let  $\mathcal{F}$  consist of the maximal dyadic cubes  $R \in \mathcal{D}_0$  for which

$$\sigma(R) > \frac{B_1}{\epsilon_0} \mu(R)$$

or

$$\sigma(R) < \delta \mu(R),$$

where  $\delta := \eta/16 = \frac{1}{32} B_1^{-1}$ . Let  $\mathcal{F}_1$  be the collection of maximal cubes  $R \in \mathcal{D}_0$  satisfying the first condition, and define  $\mathcal{F}_2$  analogously. Note that

$$\mu\left(\bigcup_{R \in \mathcal{F}_1} R\right) \leq \epsilon_0 \mu(Q),$$

so that we have by assumption (4) that

$$\sigma\left(\bigcup_{R \in \mathcal{F}_1} R\right) \leq \frac{1}{32 B_1} \sigma(Q) = \delta \sigma(Q).$$

Finally, we record that

$$\sigma\left(\bigcup_{R \in \mathcal{F}_2} R\right) = \sum_{R \in \mathcal{F}_2} \sigma(R) \leq \delta \sum_{R \in \mathcal{F}_2} \mu(R) = \delta \mu\left(\bigcup_{R \in \mathcal{F}_2} R\right) \leq \delta \mu(Q) \leq \delta \sigma(Q).$$

We may conclude that the set

$$H_1 = \bigcup_{R \in \mathcal{F}} R$$

satisfies  $\sigma(H_1) \leq 2\delta \sigma(Q) = \frac{\eta}{8} \sigma(Q)$ .

We now record the important property of the exceptional set  $H_1$ . Let  $x \in Q \setminus H_1$ . For any  $R \in \mathcal{D}_0$  satisfying that  $x \in R$  we have that

$$\frac{1}{32 B_1} = \delta \leq \frac{\sigma(R)}{\mu(R)} \leq \frac{B_1}{\epsilon_0}.$$

From this we can conclude (using a dyadic variant of Lemma 2.13 of [20]) that for all Borel sets  $A \subset \mathbb{R}^n$  there holds that

$$\delta \mu(A \cap (Q \setminus H_1)) \leq \sigma(A \cap (Q \setminus H_1)) \leq \frac{B_1}{\epsilon_0} \mu(A \cap (Q \setminus H_1)).$$

In particular, we have that  $\sigma \ll \mu$  on  $(Q \setminus H_1)$ . Using Radon–Nikodym theorem we let  $\varphi \geq 0$  be a function so that

$$\sigma(A) = \int_A \varphi d\mu$$

for all Borel sets  $A \subset Q \setminus H_1$ . We obviously have that  $\varphi \sim 1$  for  $\mu$ -a.e.  $x \in Q \setminus H_1$ .

We need another exceptional set  $H_2$ . To this end, let

$$p(x) := M_{R,m} \nu(x) = \sup_{r>0} \frac{\sigma(B(x,r))}{r^m}.$$

For  $p_0 > 0$  let  $E_{p_0} = \{p \geq p_0\}$ . Using that  $M_{R,m} : M(\mathbb{R}^n) \rightarrow L^{1,\infty}(\mu)$  boundedly we see that

$$\mu(E_{p_0}) = \mu(\{M_{R,m} \nu \geq p_0\}) \leq \frac{C}{p_0} \|\nu\| \leq \frac{C B_1}{p_0} \mu(Q).$$

We fix  $p_0 \lesssim 1$  so large that  $\mu(E_{p_0/2^m}) \leq \epsilon_0 \mu(Q)$ , so that in particular  $\sigma(E_{p_0/2^m}) \leq \frac{\eta}{8} \sigma(Q)$ . For  $x \in \{p > p_0\}$  define

$$r(x) = \sup\{r > 0 : \sigma(B(x, r)) > p_0 r^m\},$$

and then set

$$H_2 := \bigcup_{x \in \{p > p_0\}} B(x, r(x)).$$

It is clear that every ball  $B_r$  with  $\sigma(B_r) > p_0 r^m$  satisfies  $B_r \subset H_2$ . Notice that if  $y \in H_2$ , then there is  $x \in \{p > p_0\}$  so that  $y \in B(x, r(x))$ , and so  $\sigma(B(y, 2r(x))) \geq \sigma(B(x, r(x))) \geq p_0 r(x)^m = p_0 2^{-m} [2r(x)]^m$ . We conclude that  $H_2 \subset E_{p_0/2^m}$ , and so  $\sigma(H_2) \leq \frac{\eta}{8} \sigma(Q)$ .

The assumption about the set  $U_Q$  reads  $\sigma(U_Q) \leq \frac{\eta}{8} \sigma(Q)$ . Define now  $H = H_1 \cup H_2 \cup U_Q$ . The properties of  $H$  are as follows:

- (1) We have  $\sigma(H) \leq \frac{\eta}{2} \sigma(Q)$ , and so  $\sigma(H \cup T_w) \leq (1 - \eta/2) \sigma(Q) = \tau_1 \sigma(Q)$ ,  $\tau_1 < 1$ .
- (2) If  $\sigma(B_r) > p_0 r^m$ , then  $B_r \subset H$ .
- (3) We have a function  $\varphi$  so that

$$\sigma(A) = \int_A \varphi d\mu$$

for all Borel sets  $A \subset Q \setminus H$ , and  $\varphi \sim 1$  for  $\mu$ -a.e.  $x \in Q \setminus H$ .

We also have for every  $\lambda > 0$  that

$$\begin{aligned} \lambda^s \sigma(\{x \in Q \setminus H : V_{\sigma, Q} b(x) > \lambda\}) \\ &= \lambda^s \sigma(\{x \in Q \setminus H : V_Q \nu(x) > \lambda\}) \\ &\lesssim \lambda^s \mu(\{x \in Q \setminus U_Q : V_Q \nu(x) > \lambda\}) \leq B_2 \|\nu\| = B_2 \sigma(Q). \end{aligned}$$

Appealing to Theorem A.1 with the measure  $\sigma$  and the  $L^\infty$  function  $b$  we find  $G_Q \subset Q \setminus H \subset Q \setminus U_Q$  so that  $\sigma(G_Q) \gtrsim \sigma(Q)$  and

$$(2.4) \quad \|1_{G_Q} V_{\sigma, Q} f\|_{L^2(\sigma)} \lesssim \|f\|_{L^2(\sigma)}$$

for every  $f \in L^2(\sigma)$ .

Suppose now that  $g \in L^2(\mu)$  and  $\text{spt } g \subset G_Q$ . We apply Equation (2.4) with  $f = g/\varphi$  (since  $G_Q \subset Q \setminus H$  we have  $\varphi \sim 1$   $\mu$ -a.e. on the support of  $g$ ). Notice that

$$\|1_{G_Q} V_{\sigma, Q}(g/\varphi)\|_{L^2(\sigma)} = \|1_{G_Q} V_{\mu, Q} g\|_{L^2(\sigma)} \gtrsim \|1_{G_Q} V_{\mu, Q} g\|_{L^2(\mu)}$$

so that

$$\|1_{G_Q} V_{\mu, Q} g\|_{L^2(\mu)} \lesssim \|g/\varphi\|_{L^2(\sigma)} \lesssim \|g\|_{L^2(\mu)}.$$

Applying Lemma 2.1 we conclude that

$$\|1_{G_Q} V_{\mu} f\|_{L^2(\mu)} \lesssim \|f\|_{L^2(\mu)}$$

for every  $f \in L^2(\mu)$  satisfying that  $\text{spt } f \subset G_Q$ . Moreover, we have that

$$\mu(Q) \leq \sigma(Q) \lesssim \sigma(G_Q) = \int_{G_Q} \varphi d\mu \lesssim \mu(G_Q).$$

We are done. □

Let us now record the non-homogenous good lambda method of Tolsa. This is essentially Theorem 2.22 in [26], where it is done for Caldéron-Zygmund operators and without the requirement that the cubes in the statement should have small boundaries. The modified version with small boundaries is recorded in [17]. The adaptation to the square function setting is easy, and we omit it.

**2.5. Theorem.** *Let  $\mu$  be a measure of order  $m$  in  $\mathbb{R}^n$ . Let  $\beta > 0$  and  $C_1 > 0$  be big enough numbers, depending only on the dimension  $n$ , and assume  $\theta \in (0, 1)$ . Suppose for each  $(2, \beta)$ -doubling cube  $Q$  with  $C_1$ -small boundary there exists a subset  $G_Q \subset Q$  such that  $\mu(G_Q) \geq \theta\mu(Q)$  and  $V: M(\mathbb{R}^n) \rightarrow L^{1,\infty}(\mu|_{G_Q})$  is bounded with a uniform constant independent of  $Q$ . Then  $V_\mu$  is bounded in  $L^p(\mu)$  for all  $1 < p < \infty$  with a constant depending on  $p$  and on the preceding constants.*

**2.6. Remark.** In Theorem 2.5 the assumption  $V: M(\mathbb{R}^n) \rightarrow L^{1,\infty}(\mu|_{G_Q})$  can be replaced by  $V_\mu: L^2(\mu|_{G_Q}) \rightarrow L^2(\mu|_{G_Q})$ . Indeed, the latter assumption implies the former one. The original reference is Nazarov–Treil–Volberg [23], but see also Theorem 2.16 in [26].

We are ready to prove our main theorem for SFs.

*Proof of Theorem 1.6.* Proposition 2.3 gives for every  $(2, \beta)$ -doubling cube  $Q \subset \mathbb{R}^n$  with  $C_1$ -small boundary a subset  $G_Q \subset Q$  such that  $\mu(G_Q) \gtrsim \mu(Q)$  and

$$\|1_{G_Q} V_\mu f\|_{L^2(\mu)} \lesssim \|f\|_{L^2(\mu)}$$

for every  $f \in L^2(\mu)$  with  $\text{spt } f \subset G_Q$ . Applying Theorem 2.5 and Remark 2.6 gives the result.  $\square$

### 3. COMMENTS ON THE CALDERÓN–ZYGMUND CASE

In this section we describe more carefully what kind of result can directly be obtained for Calderón–Zygmund operators (this means the same thing as SIOs here) using the method of this paper. That is, we describe the analog of our main Proposition 2.3 for Calderón–Zygmund operators. For clarity we formulate a slightly less technical statement involving testing functions rather than measures – this is the one invoked in [17]. First, let us recall some definitions.

We say that  $K: \mathbb{R}^n \times \mathbb{R}^n \setminus \{(x, y) : x = y\} \rightarrow \mathbb{C}$  is an  $m$ -dimensional Calderón–Zygmund kernel if for some  $C < \infty$  and  $\alpha \in (0, 1]$  we have that

$$|K(x, y)| \leq \frac{C}{|x - y|^m}, \quad x \neq y,$$

$$|K(x, y) - K(x', y)| \leq C \frac{|x - x'|^\alpha}{|x - y|^{m+\alpha}}, \quad |x - y| \geq 2|x - x'|,$$

and

$$|K(x, y) - K(x, y')| \leq C \frac{|y - y'|^\alpha}{|x - y|^{m+\alpha}}, \quad |x - y| \geq 2|y - y'|.$$

We consider the following  $\epsilon$ -truncated singular integral operators  $T_\epsilon$ ,  $\epsilon > 0$ :

$$T_\epsilon \nu(x) = \int_{|x-y|>\epsilon} K(x, y) d\nu(y), \quad x \in \mathbb{R}^n.$$

The integral on the right hand side is absolutely convergent if, say,  $|\nu|(\mathbb{R}^n) < \infty$ .

For a positive Radon measure  $\mu$  in  $\mathbb{R}^n$  and  $f \in L^1_{\text{loc}}(\mu)$  we define

$$T_{\mu,\epsilon}f(x) = T_\epsilon(f\mu)(x), \quad x \in \mathbb{R}^n.$$

The integral defining  $T_{\mu,\epsilon}f(x)$  is absolutely convergent if for example  $f \in L^p(\mu)$  for some  $1 \leq p < \infty$  and  $\mu$  is of order  $m$ . We say that  $T_\mu$  is bounded in  $L^p(\mu)$  if the operators  $T_{\mu,\epsilon}$  are bounded in  $L^p(\mu)$  uniformly in  $\epsilon > 0$ .

We define the maximal operator  $T_*$  by

$$T_*\nu(x) = \sup_{\epsilon > 0} |T_\epsilon\nu(x)|, \quad \nu \in M(\mathbb{R}^n), x \in \mathbb{R}^n,$$

Like above, we also set  $T_{\mu,*}f(x) = T_*(f\mu)$ .

For  $p \in [1, \infty)$  we say that a function  $b_Q$  is an  $L^p(\mu)$ -admissible test function on a cube  $Q \subset \mathbb{R}^n$  (with constant  $B_1$ ), if

- (1)  $\text{spt } b_Q \subset Q$ ,
- (2)  $\mu(Q) = \int_Q b_Q d\mu$  and
- (3)  $\left(\frac{1}{\mu(Q)} \int_Q |b_Q|^p d\mu\right)^{1/p} \leq B_1$ .

Below we will need  $p > 1$  (see the comments after Theorem 1.6).

It is easier to prove local  $Tb$  theorems assuming conditions for maximal truncations  $T_{\mu,*}b_Q$  rather than uniform conditions on  $T_{\mu,\epsilon}b_Q$ . Of course, this distinction does not manifest itself in the square function setting. There is a tradeoff here and the theorems are not strictly comparable. This is because one needs much weaker conditions on  $T_{\mu,*}b_Q$  compared to  $T_{\mu,\epsilon}b_Q$ , but of course  $T_{\mu,*}b_Q$  is a larger object to begin with. A theorem involving  $T_{\mu,*}b_Q$  with very weak testing assumptions is very important also because the most efficient strategies for proving local  $Tb$  theorems involving  $T_{\mu,\epsilon}b_Q$  are based on these. See Hytönen–Nazarov [12] for some related results in the Lebesgue situation.

We now state the analog of Proposition 2.3 – the result concerning Hofmann’s problem [17] is reduced to this statement using a fancy version of Cotlar’s inequality (the main technical tool of [17]). Even the flexibility with the exceptional set  $U_Q$  is needed there.

**3.1. Proposition.** *Let  $\mu$  be a measure of degree  $m$  on  $\mathbb{R}^n$  and  $K$  be an  $m$ -dimensional kernel satisfying  $K(x, y) = -K(y, x)$ . Let  $Q \subset \mathbb{R}^n$  be a fixed cube,  $q \in (1, \infty)$  and  $b_Q$  be an  $L^q(\mu)$ -admissible test function in  $Q$  with constant  $B_1$ . Then there exists a small constant  $c_1 = c_1(q, B_1) > 0$  with the following property. If there exist  $s > 0$  and an exceptional set  $U_Q \subset \mathbb{R}^n$  so that  $\int_{U_Q} |b_Q| d\mu \leq c_1 \int_Q |b_Q| d\mu$  and*

$$\sup_{\lambda > 0} \lambda^s \mu(\{x \in Q \setminus U_Q : T_{\mu,*}b_Q(x) > \lambda\}) \leq B_2 \mu(Q) \text{ for some } B_2 < \infty,$$

*then there exists  $G_Q \subset Q \setminus E_Q$  so that  $\mu(G_Q) \gtrsim \mu(Q)$  and  $T_{\mu|G_Q} : L^2(\mu|G_Q) \rightarrow L^2(\mu|G_Q)$  with a norm depending on the constants in the assumptions.*

The proof is essentially the same than that of Proposition 2.3. However, this time we of course need to use the Calderón–Zygmund version of the big piece  $Tb$  theorem due to Nazarov–Treil–Volberg [24]. This is also proved in detail in [28]. See also [26], Theorem 5.1, for an exposition in the case of the Cauchy operator. Notice that the big piece  $Tb$  involves maximal truncations as well, which is an explanation why they appear here also.

The related local  $Tb$  corollary follows using the good lambda method. Again, we could even use testing measures, exceptional sets and so forth, but we prefer to state the slightly less technical statement here.

**3.2. Corollary.** *Let  $\mu$  be a measure of degree  $m$  on  $\mathbb{R}^n$  and  $K$  be an  $m$ -dimensional kernel satisfying  $K(x, y) = -K(y, x)$ . Suppose  $q \in (1, \infty)$ , and let  $b$  and  $t$  be large enough constants (depending only on  $n$ ). We assume that to every  $(5, b)$ -doubling cube  $Q \subset \mathbb{R}^n$  with  $t$ -small boundary there is associated an  $L^q(\mu)$ -admissible test function  $b_Q$  in  $Q$  with constant  $B_1$  such that*

$$\sup_{\lambda > 0} \lambda^s \mu(\{x \in Q: T_{\mu,*} b_Q(x) > \lambda\}) \leq B_2 \mu(Q) \text{ for some } B_2 < \infty \text{ and } s > 0.$$

*Then  $T_\mu: L^2(\mu) \rightarrow L^2(\mu)$  with a bound depending on the above constants.*

#### APPENDIX A. BIG PIECES GLOBAL $Tb$ FOR SQUARE FUNCTIONS

In this appendix we prove the big pieces global  $Tb$  theorem for square functions which we needed above. For antisymmetric Calderón–Zygmund operators with some assumptions about the maximal truncation  $T_*b$  this is by Nazarov–Treil–Volberg [24] (see also [26] and [28]). Our efficient proof in the square function setting is much more approachable than the proof in the Calderón–Zygmund case, which is why we present it here.

**A.1. Theorem.** *Let  $Q \subset \mathbb{R}^n$  be a cube. Let  $\sigma$  be a finite Borel measure in  $\mathbb{R}^n$  so that  $\text{spt } \sigma \subset Q$ . Suppose  $b$  is a function satisfying that  $\|b\|_{L^\infty(\sigma)} \leq C_b$ . For every  $w$  let  $T_w$  be the union of the maximal dyadic cubes  $R \in \mathcal{D}(w)$  for which*

$$\left| \int_R b d\sigma \right| < c_{\text{acc}} \sigma(R).$$

*We are also given a measurable set  $H \subset \mathbb{R}^n$  satisfying the following properties.*

- *There is  $\delta_0 < 1$  so that  $\sigma(H \cup T_w) \leq \delta_0 \sigma(Q)$  for every  $w$ .*
- *Every ball  $B_r$  of radius  $r$  satisfying  $\sigma(B_r) > C_0 r^m$  satisfies  $B_r \subset H$ .*
- *We have for some  $s > 0$  the estimate*

$$\sup_{\lambda > 0} \lambda^s \sigma(\{x \in Q \setminus H: V_{\sigma,Q} b(x) > \lambda\}) \leq C_1 \sigma(Q).$$

*Then there is a measurable set  $G_Q$  satisfying  $G_Q \subset Q \setminus H$  and the following properties:*

- (a)  $\sigma(G_Q) \gtrsim \sigma(Q)$ .
- (b)  $\|1_{G_Q} V_{\sigma,Q} f\|_{L^2(\sigma)} \lesssim \|f\|_{L^2(\sigma)}$  for every  $f \in L^2(\sigma)$ .

**A.2. Remark.** Only the good lambda method (Theorem 2.5) and hence the main theorem (Theorem 1.6) require the  $x$ -continuity of  $s_t$  i.e. (1.4). Proposition 2.3 and the above Theorem A.1 do not require it.

*Proof of Theorem A.1.* We begin by suppressing our operator appropriately. Set

$$S_0 = \{x \in Q: V_{\sigma,Q} b(x) > \lambda_0\},$$

where  $0 < \lambda_0 \lesssim 1$  is large enough. Now simply define

$$\tilde{s}_t(x, y) = s_t(x, y) 1_{\mathbb{R}^n \setminus S_0}(x).$$



Notice that  $(\tilde{s}_t)_{t>0}$  is a measurable family of kernels satisfying (1.2) and (1.3), which is all we shall need in what follows. Now,  $\tilde{V}_{\sigma,Q}$  (and similar objects) are defined in the natural way using the kernels  $\tilde{s}_t$ . Then for any  $f$  we have

$$(A.3) \quad \tilde{V}_{\sigma,Q}f(x) = V_{\sigma,Q}f(x)1_{\mathbb{R}^n \setminus S_0}(x) = V_{\sigma,Q}f(x)1_{\mathbb{R}^n \setminus (Q \cap \{V_{\sigma,Q}b > \lambda_0\})}(x),$$

and from here we can easily read two key things about these suppressed operators. The first is that for any  $f$  we have

$$(A.4) \quad \tilde{V}_{\sigma,Q}f(x) = V_{\sigma,Q}f(x) \text{ for } x \in \mathbb{R}^n \setminus S_0,$$

and the second is that

$$(A.5) \quad \tilde{V}_{\sigma,Q}b(x) \leq \lambda_0 \text{ for every } x \in Q.$$

Finally, with a large enough choice of  $\lambda_0$  we have (for every  $w$ ) that  $\sigma(H \cup T_w \cup S_0) \leq \delta_1 \sigma(Q)$  for some  $\delta_1 < 1$ . Indeed,

$$\sigma(S_0 \setminus H) \leq \sigma(\{x \in Q \setminus H : V_{\sigma,Q}b(x) > \lambda_0\}) \leq \frac{C_1}{\lambda_0^s} \sigma(Q).$$

At this point  $\lambda_0 \lesssim 1$  can be fixed by demanding that it satisfies

$$\lambda_0^s > \frac{2C_1}{1 - \delta_0},$$

whence we conclude that

$$(A.6) \quad \sigma(H \cup T_w \cup S_0) \leq \sigma(H \cup T_w) + \sigma(S_0 \setminus H) \leq \frac{1 + \delta_0}{2} \sigma(Q) =: \delta_1 \sigma(Q), \quad \delta_1 < 1.$$

We are now done with suppressing the operator.

We will next define the set  $G_Q$ . This is done by setting

$$p_0(x) = \mathbb{P}(\{w \in \Omega : x \in Q \setminus [H \cup T_w \cup S_0]\}),$$

and then defining

$$G_Q = \left\{ x \in Q : p_0(x) > \frac{1 - \delta_1}{2} =: \tau \right\} \subset Q \setminus H.$$

An argument by Nazarov–Treil–Volberg (see [24]) shows that  $\sigma(G_Q) \gtrsim \sigma(Q)$ . Indeed, the argument goes as follows. Notice first that by (A.6) we have that

$$\int_Q p_0(x) d\sigma(x) = \int_\Omega \sigma(Q \setminus [H \cup T_w \cup S_0]) d\mathbb{P}(w) \geq (1 - \delta_1) \sigma(Q).$$

Since  $1 - p_0 \geq 0$  everywhere, and  $1 - p_0 \geq 1 - \tau = (1 + \delta_1)/2$  on  $Q \setminus G_Q$ , we have

$$\int_Q (1 - p_0(x)) d\sigma(x) \geq \int_{Q \setminus G_Q} (1 - p_0(x)) d\sigma(x) \geq \frac{1 + \delta_1}{2} \sigma(Q \setminus G_Q).$$

We conclude that

$$\sigma(Q \setminus G_Q) \leq \frac{2}{1 + \delta_1} \left( \sigma(Q) - \int_Q p_0(x) d\sigma(x) \right) \leq \frac{2\delta_1}{1 + \delta_1} \sigma(Q),$$

and so

$$\sigma(G_Q) \geq \left( 1 - \frac{2\delta_1}{1 + \delta_1} \right) \sigma(Q) = \frac{1 - \delta_1}{1 + \delta_1} \sigma(Q).$$

It remains to prove that  $\|1_{G_Q} V_{\sigma,Q} f\|_{L^2(\sigma)} \lesssim \|f\|_{L^2(\sigma)}$  for every  $f \in L^2(\sigma)$ . The key property of  $G_Q$  is as follows. Suppose  $h \geq 0$  is any positive function. Then we have that

$$\int_{G_Q} h(x) d\sigma(x) \leq \tau^{-1} \int_{G_Q} p_0(x) h(x) d\sigma(x) = \tau^{-1} E_w \int_{G_Q \setminus [H \cup T_w \cup S]} h(x) d\sigma(x).$$

We apply this as follows:

$$\begin{aligned} \|1_{G_Q} V_{\sigma,Q} f\|_{L^2(\sigma)}^2 &= \int_{G_Q} \int_0^{\ell(Q)} |\theta_t^\sigma f(x)|^2 \frac{dt}{t} d\sigma(x) \\ &\leq \tau^{-1} E_w \int_{G_Q \setminus [H \cup T_w \cup S_0]} \int_0^{\ell(Q)} |\theta_t^\sigma f(x)|^2 \frac{dt}{t} d\sigma(x) \\ &= \tau^{-1} E_w \sum_{R \in \mathcal{D}_0} \int_{[R \cap G_Q] \setminus [H \cup T_w \cup S_0]} \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\theta_t^\sigma f(x)|^2 \frac{dt}{t} d\sigma(x), \end{aligned}$$

where again  $\mathcal{D}_0 = \mathcal{D}(0)$ . Given  $w$  we then write

$$\sum_{R \in \mathcal{D}_0} = \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-good}}} + \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-bad}}},$$

where  $R \in \mathcal{D}_0$  is said to be  $\mathcal{D}(w)$ -good if  $d(R, \partial P) > \ell(R)^\gamma \ell(P)^{1-\gamma}$  for every  $P \in \mathcal{D}(w)$  satisfying  $\ell(P) \geq 2^r \ell(R)$ . Here  $r \gtrsim 1$  is a fixed large enough parameter, and  $\gamma := \alpha/(2m+2\alpha)$ . It is a standard fact by Nazarov–Treil–Volberg (see [25]) that given  $R \in \mathcal{D}_0$  we have that

$$(A.7) \quad \mathbb{P}(\{w \in \Omega : R \text{ is } \mathcal{D}(w)\text{-bad}\}) \leq \tau/2$$

for a large enough fixed  $r$ .

Using (A.7) we estimate

$$\begin{aligned} E_w \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-bad}}} \int_{[R \cap G_Q] \setminus [H \cup T_w \cup S_0]} \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\theta_t^\sigma f(x)|^2 \frac{dt}{t} d\sigma(x) \\ \leq \sum_{R \in \mathcal{D}_0} \mathbb{P}(\{w \in \Omega : R \text{ is } \mathcal{D}(w)\text{-bad}\}) \int_{R \cap G_Q} \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\theta_t^\sigma f(x)|^2 \frac{dt}{t} d\sigma(x) \\ \leq \frac{\tau}{2} \int_{G_Q} \int_0^{\ell(Q)} |\theta_t^\sigma f(x)|^2 \frac{dt}{t} d\sigma(x). \end{aligned}$$

To be precise, for the following we would need the a priori finiteness of this term. However, this is easy to arrange in a multiple of ways, so we skip this technicality. We may

now conclude (using also that  $\theta_t^\sigma f(x) = \tilde{\theta}_t^\sigma f(x)$  for every  $x \in Q \setminus S_0$  by (A.4)) that

$$\begin{aligned} & \|1_{G_Q} V_{\sigma, Q} f\|_{L^2(\sigma)}^2 \\ & \leq 2\tau^{-1} E_w \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-good}}} \int_{[R \cap G_Q] \setminus [H \cup T_w \cup S_0]} \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\tilde{\theta}_t^\sigma f(x)|^2 \frac{dt}{t} d\sigma(x) \\ & \lesssim E_w \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-good} \\ R \not\subset H \cup T_w}} \int_R \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\tilde{\theta}_t^\sigma f(x)|^2 \frac{dt}{t} d\sigma(x). \end{aligned}$$

We will now fix  $w$ , write  $\mathcal{D} = \mathcal{D}(w)$  and  $T = T_w$ , and prove that

$$(A.8) \quad \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}\text{-good} \\ R \not\subset H \cup T}} \int_R \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\tilde{\theta}_t^\sigma f(x)|^2 \frac{dt}{t} d\sigma(x) \lesssim \|f\|_{L^2(\sigma)}^2.$$

This will then end the proof.

The important property of the set  $T$  is that if  $R \in \mathcal{D}$  and  $R \not\subset T$  then

$$\left| \int_R b d\sigma \right| \gtrsim \sigma(R),$$

while the important property of the set  $H$  is that if  $L \subset \mathbb{R}^n$  is an arbitrary cube satisfying  $L \not\subset H$  then  $\sigma(\lambda L) \lesssim \lambda^m \ell(L)^m$  for all  $\lambda \geq 1$ . It is useful to say that  $R \in \mathcal{D}_0^{tr}$  (tr stands for *transit*) if  $R \in \mathcal{D}_0$ ,  $\sigma(R) \neq 0$  and  $R \not\subset H \cup T$ , and  $P \in \mathcal{D}^{tr}$  if  $P \in \mathcal{D}$ ,  $\sigma(P) \neq 0$  and  $P \not\subset H \cup T$ . Note that  $\mathcal{D}_0^{tr}$  really means  $w$ -transit cubes from  $\mathcal{D}_0$  (and one should really write  $\mathcal{D}_0^{tr}(w)$ ), but  $w$  is fixed and so  $T$  is fixed and we do not need to insist on this.

It is time to expand the function  $f$  in the grid  $\mathcal{D}$  using  $b$ -adapted martingales only in the transit cubes  $P \in \mathcal{D}^{tr}$ . Denote  $\langle f \rangle_A = \langle f \rangle_A^\sigma = \sigma(A)^{-1} \int_A f d\sigma$ , if  $\sigma(A) \neq 0$ . Let  $P_0 = Q^*(w)$  (see the Definition 2.2) so that all  $P \in \mathcal{D}$  satisfy  $P \subset P_0$ . Without loss of generality we can assume that  $\text{spt } b \subset Q$  and  $\text{spt } f \subset Q$ . Define

$$E_{P_0} f = \frac{\langle f \rangle_{P_0}}{\langle b \rangle_{P_0}} b.$$

(This is actually independent of  $w$  since it just equals  $E_Q f$ , because  $\text{spt } \sigma \subset Q \subset P_0$ ). For any cube  $P \in \mathcal{D}^{tr}$  define the function  $\Delta_P f$  as follows:

$$\Delta_P f = \sum_{\substack{P' \in \text{ch}(P) \\ \sigma(P') \neq 0}} A_{P'}(f) 1_{P'},$$

where

$$A_{P'}(f) = \begin{cases} \left( \frac{\langle f \rangle_{P'}}{\langle b \rangle_{P'}} - \frac{\langle f \rangle_P}{\langle b \rangle_P} \right) b & \text{if } P' \in \mathcal{D}^{tr}, \\ f - \frac{\langle f \rangle_P}{\langle b \rangle_P} b & \text{if } P' \notin \mathcal{D}^{tr}. \end{cases}$$

Notice that  $P_0 \in \mathcal{D}^{tr}$ , since  $\sigma(P_0) = \sigma(Q)$  and every non-transit cube  $P$  has to satisfy  $\sigma(P) \leq \sigma(H \cup T) \leq \delta_0 \sigma(Q)$ . It is easy to see that

$$f = \sum_{P \in \mathcal{D}^{tr}} \Delta_P f + E_{P_0} f$$

$\sigma$ -a.e. and in  $L^2(\sigma)$ , and that

$$\sum_{P \in \mathcal{D}^{tr}} \|\Delta_P f\|_{L^2(\sigma)}^2 + \|E_{P_0} f\|_{L^2(\sigma)}^2 \lesssim \|f\|_{L^2(\sigma)}^2.$$

See e.g. Section 5.4.4 of [26]. It will be convenient to exploit notation by redefining on the largest level  $P_0$  the operator  $\Delta_{P_0} f$  to be  $\Delta_{P_0} f + E_{P_0} f$ .

Going back to (A.8) we see that we need to control

$$\sum_{\substack{R \in \mathcal{D}_0^{tr} \\ R \text{ is } \mathcal{D}\text{-good}}} \int_R \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} \left| \sum_{P \in \mathcal{D}^{tr}} \tilde{\theta}_t^\sigma \Delta_P f(x) \right|^2 \frac{dt}{t} d\sigma(x).$$

Given  $R \in \mathcal{D}_0^{tr}$ ,  $R$  is  $\mathcal{D}(w)$ -good, the  $P \in \mathcal{D}^{tr}$  summation is split in to the following four pieces:

- (1)  $P: \ell(P) < \ell(R)$ ;
- (2)  $P: \ell(P) \geq \ell(R)$  and  $d(P, R) > \ell(R)^\gamma \ell(P)^{1-\gamma}$ ;
- (3)  $P: \ell(R) \leq \ell(P) \leq 2^r \ell(R)$  and  $d(P, R) \leq \ell(R)^\gamma \ell(P)^{1-\gamma}$ ;
- (4)  $P: \ell(P) > 2^r \ell(R)$  and  $d(P, R) \leq \ell(R)^\gamma \ell(P)^{1-\gamma}$ .

For future need we set

$$A_{PR} := \frac{\ell(P)^{\alpha/2} \ell(R)^{\alpha/2}}{D(P, R)^{m+\alpha}} \sigma(P)^{1/2} \sigma(R)^{1/2};$$

$$D(P, R) := \ell(P) + \ell(R) + d(P, R).$$

The following estimate by Nazarov–Treil–Volberg (see e.g. [25]) is extremely useful

$$\sum_{\substack{P \in \mathcal{D}^{tr} \\ R \in \mathcal{D}_0^{tr}}} A_{PR} x_P y_R \lesssim \left( \sum_P x_P^2 \right)^{1/2} \left( \sum_R y_R^2 \right)^{1/2}$$

for every  $x_P, y_R \geq 0$ . For an easy reference, see pp. 159–160 in [26]. In particular, we have that

$$\left( \sum_{R \in \mathcal{D}_0^{tr}} \left[ \sum_{P \in \mathcal{D}^{tr}} A_{PR} x_P \right]^2 \right)^{1/2} \lesssim \left( \sum_P x_P^2 \right)^{1/2}.$$

The sums (1) and (2) are handled as follows. Notice that in (1) we have  $\ell(P) < \ell(R) \leq \ell(P_0)$  so that  $\int \Delta_P f d\sigma = 0$ . Therefore, using the  $y$ -Hölder for  $\tilde{s}_t$  we get

$$(A.9) \quad |\tilde{\theta}_t^\sigma \Delta_P f(x)| \lesssim A_{PR} \sigma(R)^{-1/2} \|\Delta_P f\|_{L^2(\sigma)}, \quad (x, t) \in W_R,$$

where  $W_R := R \times [\ell(R)/2, \ell(R))$ . In the case (2), the size estimate for  $\tilde{s}_t$  yields

$$|\tilde{\theta}_t^\sigma \Delta_P f(x)| \lesssim \frac{\ell(R)^\alpha}{d(P, R)^{m+\alpha}} \sigma(P)^{1/2} \|\Delta_P f\|_{L^2(\sigma)}, \quad (x, t) \in W_R.$$

But this yields the same bound as in (A.9), since here

$$\frac{\ell(R)^\alpha}{d(P, R)^{m+\alpha}} \sigma(P)^{1/2} \lesssim A_{PR} \sigma(R)^{-1/2}.$$

To see this, notice that it is obvious if  $d(P, R) \geq \ell(P)$ . In the opposite case note that  $d(P, R)^{m+\alpha} \gtrsim D(P, R)^{m+\alpha} \ell(P)^{-\alpha/2} \ell(R)^{\alpha/2}$ . This is seen by combining the facts that  $d(P, R) > \ell(R)^\gamma \ell(P)^{1-\gamma}$ ,  $\gamma m + \gamma \alpha = \alpha/2$  and  $D(P, R) \lesssim \ell(P)$ . Thus, also in the case (2) the estimate (A.9) holds. The cases (1) and (2) are therefore under control via the estimate

$$\sum_{R \in \mathcal{D}_0^{tr}} \left[ \sum_{P \in \mathcal{D}^{tr}} A_{PR} \|\Delta_P f\|_{L^2(\sigma)} \right]^2 \lesssim \sum_{P \in \mathcal{D}^{tr}} \|\Delta_P f\|_{L^2(\sigma)}^2 \lesssim \|f\|_{L^2(\sigma)}^2.$$

The summation (3) is even easier. Using that  $P$  and  $R$  are both transit,  $t \sim \ell(R) \sim \ell(P)$  and the size estimate for  $\tilde{s}_t$  we see that

$$|\tilde{\theta}_t^\sigma \Delta_P f(x)| \lesssim t^{-m} \sigma(P)^{1/2} \|\Delta_P f\|_{L^2(\sigma)} \lesssim \sigma(R)^{-1/2} \|\Delta_P f\|_{L^2(\sigma)}, \quad (x, t) \in W_R.$$

This can then easily be summed, since given  $R$  there are only finitely many  $P$  such that  $\ell(P) \sim \ell(R)$  and  $d(P, R) \lesssim \min(\ell(P), \ell(R))$ .

We move on to the main term (4). For each  $R \in \mathcal{D}_0^{tr}$  satisfying that  $R$  is  $\mathcal{D}$ -good,  $R \subset P_0$  and  $\ell(R) < 2^{-r} \ell(P_0)$  we let  $P_{R,k} \in \mathcal{D}$ ,  $k \in \{r, r+1, \dots, \log_2[\ell(P_0)/\ell(R)]\}$ , be the unique  $\mathcal{D}$ -cube satisfying that  $\ell(P_{R,k}) = 2^k \ell(R)$  and  $R \subset P_{R,k}$ . Such a cube exists since  $R$  is  $\mathcal{D}$ -good. Moreover, since  $R \not\subset H \cup T$  then also  $P_{R,k} \not\subset H \cup T$  i.e.  $P_{R,k} \in \mathcal{D}^{tr}$ . We see that we only need to prove that

$$\sum_{\substack{R \in \mathcal{D}_0^{tr}: R \subset P_0 \\ R \text{ is } \mathcal{D}\text{-good} \\ \ell(R) < 2^{-r} \ell(P_0)}} \int_R \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} \left| \sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} \tilde{\theta}_t^\sigma \Delta_{P_{R,k}} f(x) \right|^2 \frac{dt}{t} d\sigma(x) \lesssim \|f\|_{L^2(\sigma)}^2.$$

Recalling that all  $P_{R,k}$ ,  $r \leq k \leq \log_2[\ell(P_0)/\ell(R)]$  are transit, we see using a standard calculation that  $\sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} \tilde{\theta}_t^\sigma \Delta_{P_{R,k}} f$  equals

$$\begin{aligned} & - \sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} B_{P_{R,k-1}} \tilde{\theta}_t^\sigma (1_{\mathbb{R}^n \setminus P_{R,k-1}} b) \\ & + \sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} \tilde{\theta}_t^\sigma (1_{P_{R,k} \setminus P_{R,k-1}} \Delta_{P_{R,k}} f) + \frac{\langle f \rangle_{P_{R,r}}}{\langle b \rangle_{P_{R,r}}} \tilde{\theta}_t^\sigma b, \end{aligned}$$

where

$$B_{P_{R,k-1}} = \langle \Delta_{P_{R,k}} f / b \rangle_{P_{R,k-1}} = \begin{cases} \frac{\langle f \rangle_{P_{R,k-1}}}{\langle b \rangle_{P_{R,k-1}}} - \frac{\langle f \rangle_{P_{R,k}}}{\langle b \rangle_{P_{R,k}}}, & \text{if } r+1 \leq k < \log_2 \frac{\ell(P_0)}{\ell(R)}, \\ \frac{\langle f \rangle_{P_{R,k-1}}}{\langle b \rangle_{P_{R,k-1}}}, & k = \log_2 \frac{\ell(P_0)}{\ell(R)}. \end{cases}$$

Let us start deciphering this by proving that the term

$$\Pi := \sum_{\substack{R \in \mathcal{D}_0^{tr}: R \subset P_0 \\ R \text{ is } \mathcal{D}\text{-good} \\ \ell(R) < 2^{-r} \ell(P_0)}} \left| \frac{\langle f \rangle_{P_{R,r}}}{\langle b \rangle_{P_{R,r}}} \right|^2 \int_R \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\tilde{\theta}_t^\sigma b(x)|^2 \frac{dt}{t} d\sigma(x)$$

is under control. We simply estimate

$$\Pi \lesssim \sum_{P \in \mathcal{D}^{tr}} |\langle f \rangle_P|^2 a_P, \quad a_P := \sum_{\substack{R \in \mathcal{D}_0^{tr}: R \subset P_0 \\ R \text{ is } \mathcal{D}\text{-good} \\ \ell(R) < 2^{-r} \ell(P_0) \\ P_{R,r} = P}} \int_R \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\tilde{\theta}_t^\sigma b(x)|^2 \frac{dt}{t} d\sigma(x).$$

To have  $\Pi \lesssim \|f\|_{L^2(\sigma)}^2$  it is enough to verify the Carleson property of  $(a_P)_{P \in \mathcal{D}}$ . To this end, let  $S \in \mathcal{D}$  be arbitrary. We have that

$$\begin{aligned} \sum_{\substack{P \in \mathcal{D} \\ P \subset S}} a_P &\leq \sum_{\substack{R \in \mathcal{D}_0^{tr} \\ R \subset S}} \iint_{[S \times (0, \ell(Q))] \cap W_R} |\tilde{\theta}_t^\sigma b(x)|^2 \frac{dt}{t} d\sigma(x) \\ &\leq \iint_{S \times (0, \ell(Q))} |\tilde{\theta}_t^\sigma b(x)|^2 \frac{dt}{t} d\sigma(x) = \int_S [\tilde{V}_{\sigma, Q} b(x)]^2 d\sigma(x) \lesssim \sigma(S), \end{aligned}$$

since  $\tilde{V}_{\sigma, Q} b(x) \lesssim 1$  for every  $x \in \text{spt } \sigma$  by (A.5).

We are only left with some completely standard calculations (but we need to be slightly careful to use transitivity). So let us first control  $|B_{P_{R,k-1}} \tilde{\theta}_t^\sigma(1_{\mathbb{R}^n \setminus P_{R,k-1}} b)(x)|$  for  $(x, t) \in W_R$ . Notice that  $R \subset B(x, d(R, \partial P_{R,k-1})/2)$ , since  $d(R, \partial P_{R,k-1}) \geq 2^{r(1-\gamma)} \ell(R) \geq C_d \ell(R)$  by having  $r$  large enough to begin with. The point is that  $B(x, d(R, \partial P_{R,k-1})/2) \not\subset H$ . Moreover, we clearly have that  $B(x, d(R, \partial P_{R,k-1})/2) \subset P_{R,k-1}$ . Using these facts we get

$$\begin{aligned} |\tilde{\theta}_t^\sigma(1_{\mathbb{R}^n \setminus P_{R,k-1}} b)(x)| &\lesssim \int_{\mathbb{R}^n \setminus B(x, d(R, \partial P_{R,k-1})/2)} \frac{\ell(R)^\alpha}{|x-y|^{m+\alpha}} d\sigma(y) \\ &\lesssim \ell(R)^\alpha d(R, \partial P_{R,k-1})^{-\alpha} \lesssim \left( \frac{\ell(R)}{\ell(P_{R,k-1})} \right)^{\alpha/2} \sim 2^{-\alpha k/2}, \end{aligned}$$

where we also used that  $d(R, \partial P_{R,k-1}) \geq \ell(R)^{1/2} \ell(P_{R,k-1})^{1/2}$  (which follows since  $R$  is  $\mathcal{D}$ -good). Since  $P_{R,k-1} \not\subset T$  we have

$$\begin{aligned} |B_{P_{R,k-1}} \sigma(P_{R,k-1})| &\lesssim \left| \int_{P_{R,k-1}} B_{P_{R,k-1}} b d\sigma \right| \\ &= \left| \int_{P_{R,k-1}} \Delta_{P_{R,k}} f d\sigma \right| \leq \sigma(P_{R,k-1})^{1/2} \|\Delta_{P_{R,k}} f\|_{L^2(\sigma)}. \end{aligned}$$

Combining these estimates we get for  $(x, t) \in W_R$  that

$$(A.10) \quad |B_{P_{R,k-1}} \tilde{\theta}_t^\sigma(1_{\mathbb{R}^n \setminus P_{R,k-1}} b)(x)| \lesssim 2^{-\alpha k/2} \sigma(P_{R,k-1})^{-1/2} \|\Delta_{P_{R,k}} f\|_{L^2(\sigma)}.$$

Let us still estimate  $|\tilde{\theta}_t^\sigma(1_{P_{R,k} \setminus P_{R,k-1}} \Delta_{P_{R,k}} f)(x)|$  for  $(x, t) \in W_R$ . Let  $S \in \text{ch}(P_{R,k})$ ,  $S \neq P_{R,k-1}$ . We do not know whether this cube is transitive or not, but it shall not



matter. Indeed, we just estimate

$$\begin{aligned} |\tilde{\theta}_t^\sigma(1_S \Delta_{P_{R,k}} f)(x)| &\lesssim \frac{\ell(R)^\alpha}{d(R, S)^{m+\alpha}} \int_{P_{R,k}} |\Delta_{P_{R,k}} f(y)| d\sigma(y) \\ &\lesssim \left( \frac{\ell(R)}{\ell(P_{R,k-1})} \right)^{\alpha/2} \frac{\sigma(P_{R,k})^{1/2}}{\ell(P_{R,k-1})^m} \|\Delta_{P_{R,k}} f\|_{L^2(\sigma)} \\ &\lesssim 2^{-\alpha k/2} \sigma(P_{R,k-1})^{-1/2} \|\Delta_{P_{R,k}} f\|_{L^2(\sigma)}, \end{aligned}$$

where we used that  $\ell(S) = \ell(P_{R,k-1})$ ,  $d(R, S)^{m+\alpha} \geq \ell(R)^{\alpha/2} \ell(S)^{\alpha/2} \ell(S)^m$  and the transitivity of  $P_{R,k-1}, P_{R,k}$ . So  $|\tilde{\theta}_t^\sigma(1_{P_{R,k} \setminus P_{R,k-1}} \Delta_{P_{R,k}} f)(x)|$  satisfies the same estimate as in (A.10).

We are done with the proof if we can control the summation

$$\sum_{\substack{R \in \mathcal{D}_0^{tr}: R \subset P_0 \\ R \text{ is } \mathcal{D}\text{-good} \\ \ell(R) < 2^{-r} \ell(P_0)}} \sigma(R) \left[ \sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} 2^{-\alpha k/2} \sigma(P_{R,k-1})^{-1/2} \|\Delta_{P_{R,k}} f\|_{L^2(\sigma)} \right]^2.$$

Using a summation argument that appears in p. 9 in [15] we dominate this by

$$\sum_{P \in \mathcal{D}^{tr}} \|\Delta_P f\|_{L^2(\sigma)}^2 \lesssim \|f\|_{L^2(\sigma)}^2.$$

□

## REFERENCES

- [1] P. Auscher, S. Hofmann, C. Muscalu, T. Tao, C. Thiele, Carleson measures, trees, extrapolation, and  $T(b)$  theorems, *Publ. Mat.* 46 (2) (2002) 257–325.
- [2] P. Auscher, S. Hofmann, M. Lacey, A. McIntosh, P. Tchamitchian, The solution of the Kato square root problem for second order elliptic operators on  $\mathbb{R}^n$ , *Ann. of Math.* 156 (2) (2002) 633–654.
- [3] P. Auscher, E. Routin, Local  $Tb$  theorems and Hardy inequalities, *J. Geom. Anal.* 23 (1) (2013) 303–374.
- [4] P. Auscher, Q. X. Yang, BCR algorithm and the  $T(b)$  theorem, *Publ. Mat.* 53 (1) (2009) 179–196.
- [5] M. Christ, A  $T(b)$  theorem with remarks on analytic capacity and the Cauchy integral, *Colloq. Math.* 50/51 (1990) 601–628.
- [6] V. Chousionis, J. Garnett, T. Le, X. Tolsa, Square functions and uniform rectifiability, *Trans. Amer. Math. Soc.* 368 (2016) 6063–6102.
- [7] G. David, Unrectifiable 1-sets have vanishing analytic capacity, *Rev. Mat. Iberoam.* 14 (1998) 369–479.
- [8] G. David, J.-L. Journé, A boundedness criterion for generalized Calderón-Zygmund operators, *Ann. of Math.* 120 (1984) 371–397.
- [9] S. Hofmann, A local  $Tb$  theorem for square functions, *Perspectives in partial differential equations, harmonic analysis and applications* (Providence, RI, 2008), pp. 175–185, *Proc. Sympos. Pure Math.*, vol. 79, Amer. Math. Soc. (2008).
- [10] S. Hofmann, A proof of the local  $Tb$  theorem for standard Calderón-Zygmund operators, unpublished manuscript, arXiv:0705.0840, 2007.
- [11] T. Hytönen, The sharp weighted bound for general Calderón-Zygmund operators, *Ann. of Math.* 175 (2012) 1473–1506.
- [12] T. Hytönen, F. Nazarov, The local  $Tb$  theorem with rough test functions, preprint, arXiv:1206.0907, 2012.
- [13] M. Lacey, H. Martikainen, Local  $Tb$  theorem with  $L^2$  testing conditions and general measures: Calderón-Zygmund operators, *Ann. Sci. Éc. Norm. Supér.* 49 (2016) 57–86.
- [14] M. Lacey, H. Martikainen, Local  $Tb$  theorem with  $L^2$  testing conditions and general measures: Square functions, *J. Anal. Math.* 133 (2017) 71–89.

- [15] H. Martikainen, M. Mourgoglou, Square functions with general measures, *Proc. Amer. Math. Soc.* 142 (2014) 3923–3931.
- [16] H. Martikainen, M. Mourgoglou, Boundedness of non-homogeneous square functions and  $L^q$  type testing conditions with  $q \in (1, 2)$ , *Math. Res. Lett.* 22 (2015) 1417–1457.
- [17] H. Martikainen, M. Mourgoglou, X. Tolsa, Improved Cotlar’s inequality in the context of local  $Tb$  theorems, *J. Funct. Anal.* 274 (2018) 1255–1275.
- [18] H. Martikainen, M. Mourgoglou, E. Vuorinen, Non-homogeneous square functions on general sets: suppression and big pieces methods, *J. Geom. Anal.* 27 (2017) 3176–3227.
- [19] H. Martikainen, E. Vuorinen, Dyadic-probabilistic methods in bilinear analysis, *Mem. Amer. Math. Soc.*, to appear, arXiv:1609.01706, 2016.
- [20] P. Mattila, *Geometry of sets and measures in Euclidean spaces: Fractals and rectifiability*, Cambridge Studies in Advanced Mathematics, Vol. 44, Cambridge University Press, Cambridge, 1995.
- [21] S. Mayboroda, A. Volberg, Boundedness of the square function and rectifiability, *C. R. Math. Acad. Sci. Paris* 347 (2009) 1051–1056.
- [22] F. Nazarov, S. Treil, A. Volberg, Accretive system  $Tb$ -theorems on nonhomogeneous spaces, *Duke Math. J.* 113 (2) (2002) 259–312.
- [23] F. Nazarov, S. Treil, A. Volberg, Weak type estimates and Cotlar inequalities for Calderón-Zygmund operators on nonhomogeneous spaces, *Int. Math. Res. Not. IMRN* 1998 (1998) 463–487.
- [24] F. Nazarov, S. Treil, A. Volberg, The  $Tb$ -theorem on non-homogeneous spaces that proves a conjecture of Vitushkin, unpublished manuscript, arXiv:1401.2479.
- [25] F. Nazarov, S. Treil, A. Volberg, The  $Tb$ -theorem on non-homogeneous spaces, *Acta Math.* 190 (2) (2003) 151–239.
- [26] X. Tolsa, *Analytic capacity, the Cauchy transform, and non-homogeneous Calderón-Zygmund theory*, Progress in Mathematics, Vol. 307, Birkhäuser Verlag, Basel, 2014.
- [27] X. Tolsa, Painlevé’s problem and the semiadditivity of analytic capacity, *Acta Math.* 190 (2003) 105–149.
- [28] A. Volberg, Calderón-Zygmund capacities and operators on nonhomogeneous spaces, *CBMS Regional Conference Series in Mathematics* vol. 100 (2003) pp. 1–165.

(H.M.) DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, P.O.B. 68, FI-00014 UNIVERSITY OF HELSINKI, FINLAND

*E-mail address:* henri.martikainen@helsinki.fi

(M.M.) DEPARTAMENT DE MATEMÀTIQUES, UNIVERSITAT AUTÒNOMA DE BARCELONA AND CENTRE DE RESERCA MATEMÀTICA, EDIFICI C FACULTAT DE CIÈNCIES, 08193 BELLATERRA (BARCELONA)

*Current address:* Departamento de Matemáticas, Universidad del País Vasco, Barrio Sarriena s/n 48940 Leioa, Spain and, Ikerbasque, Basque Foundation for Science, Bilbao, Spain.

*E-mail address:* michail.mourgoglou@ehu.eus

(E.V.) DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, P.O.B. 68, FI-00014 UNIVERSITY OF HELSINKI, FINLAND

*E-mail address:* emil.vuorinen@helsinki.fi